THE IMPACTS OF SOME DEMOGRAPHIC AND HIV RELATED RISK BEHAVIOR VARIABLES ON THE STATUS OF HIV/AIDS: THE CASE OF OSSA MOBILE VCT CENTER IN URBAN ADDIS ABABA

BY

HUSSEIN MOHAMMED

DEPARTMENT OF STATISTICS

FACULTY OF SCIENCE

A THESIS SUBMITTED TO THE OFFICE OF GRADUATE PROGRAMMES OF ADDIS ABEBA UNIVERSITY, IN PARTIAL FULFILLMENT FOR THE AWARD OF MASTER OF SCIENCE IN STATISTICS

ADDIS ABABA

JULY, 2007

ACKNOWLEDGMENTS

I am very grateful to Professor **Eshetu Wencheko**, my thesis advisor and instructor for his invaluable comments and suggestions that contributed to the successful realization of the study.

Many thanks are due to my institution Oromiya Finance and Economic development Bureau (BoFED) which allowed me to join the **MSc**. Program and to utilize computer and other facilities without limitations.

I am indebted to Ato Dagne Lemma who helped me to continue the graduate study in statistics. I also extend my gratitude to the entire staff of the department of statistics for their assistance in various ways.

Finally, I would like to thank Ato Shibru Temesgen and W/rt Meseret Dagne, for their consistent support as well as my friends for their encouragement and support throughout the course of the study.

CONTENTS

Acknowledgment...i
Abstract...v
Acronym..viii

Chapter One
1. Introduction...1
1.1. Rationale for the study 1
1.2. Socio-demographic profile among VCT clients 3
1.2. Objectives of the study...5
1.3. Definition of Terms...6
1.4. Possible applications of the result..7
1.5. Limitations 7

Chapter two
2. Literature review 8
2.1. Voluntary counselling and testing for HIV.......................................8
2.2. HIV Related risk behaviors...10

Chapter Three
3. Data and Methodology 13
3.1. The Data...13
3.2. Variables of interest 13
3.3. Methodology 16
3.3.1. The Logistic Regression 16
3.3.2. The model 16
3.3.3. Parameter Estimation 19
3.3.4. Assessing the goodness of fit...20
3.3.4.1. The likelihood ratio criterion (deviance analysis) 20
3.3.4.2. Pearson's X^2- statistic..22
3.3.4.3. Hosmer-Lemeshow Test 23
3.3.4.4. Test of goodness of fit of the Final Model-Wald Test...................24

Chapter four
4. Data Analysis...26
4.1. Summary Statistics...26
4.1.1. Demographic profile 26
4.1.2. HIV related risky behavior..28
4.2. Univariate Findings...31
4.3. Multivariate Findings...34
4.3.1. Estimates for logistic model..34
4.3.1.1. Variables in the Equation...36
4.4. Assessing the Goodness of fit of the model.......................................38
4.4.1. The likelihood ratio test 38
4.4.2. Classification table...39

4.4.3. The Hosmer-Lemeshow test 40
4.5. Case Diagnostics 41
4.6. Interpretation of the logistic regression model 42

Chapter Five

5. Conclusions and Recommendations

5.1. Conclusions 44
5.2. Recommendations 45

Abstract

Voluntary counselling and testing (VCT) for HIV allows individuals to determine their HIV status and serve as a gateway for both HIV prevention and early access to treatment, care and support. The services are available in government, non-government and health facilities. However, very little is known about the demographic profile, risk behavior and HIV prevalence among VCT clients.

The study aims at describing the impacts of some pre-test counselling session factors on the status of HIV/AIDS, which includes both demographic and risk behavior factors. The data for this study were taken from the OSSA mobile VCT center in urban Addis Ababa and data were analyzed using SPSS.

Among the 14810 clients a majority (67.6%) were males with overall average age of 28.8 years. About (62%) of the total clients were never-married followed by currently married clients (29.7%) among those whose marriage was dissolved (divorced and widowed). The proportion of female clients is higher than males. About (57.1%) of the clients have attended junior and high school and female clients appeared to be less educated than males. The reason cited for seeking VCT services by the majority (49.4%) of the clients is due to the risky behavior they passed through followed by those who went to" plan for future" (44.7%). The vast majority of the clients (78.2%) ever had sexual practices in the past and female clients were less likely than their male counterparts to report sexual practices. Nearly half of (52.2%) of the clients reported that they did not use condoms in three months time prior to testing and during the last sex months; females are more likely than males for not using condoms. Prevalence of HIV infection among mobile VCT clients was estimated to be 6.4% and varied by socio-demographic and risk behavior

characteristics of the clients. The findings show that HIV status of an individual is strongly associated with independent variables included in the analysis except for employment status. The independent variables are: - age, sex, marital status, educational level, employment status, occupation, the most important reason why client is here today, previously tested, ever had sex with penetration, suspected exposure time, condom use last three months, used condom last time had sex and history of STI. Most of the variables included in the regression model had significant effects.

Considering the prevailing high level of HIV infection rate among VCT clients, appropriate behavioral change strategies need to be designed particularly focused on people with low socio-economic status and high-risk behavior clients to help them lead healthy lives. In conclusion, strong referral linkages need to be established for the provision, care and support services in urban Addis Ababa.

Acronym

UN AIDS	United Nation program on HIV/AIDS
MOH	Ministry of Health
CSA	Central Statistical Agency
VCT	Voluntary counselling and Testing
OSSA	Organization for Social Services for AIDS
PEPFAR	President's Emergency Plan for AIDS Relief
CDC	United States Center for Disease Control and Prevention
STI	Sexually Transmitted Infection
ARV	Anti Retroviral Drugs
MTCT	Mother-to-Child Transmission
FGAE	Family Guidance Association of Ethiopia

CHAPTER ONE

INTRODUCTION

1.1 Rationale for the study

HIV/AIDS is the most formidable epidemic in recorded history. The HIV/AIDS epidemic has become a serious health and development problem in many countries around the world. According to the joint United Nation programme on HIV/AIDS (UN AIDS), an estimated 39.5 million people world wide were living with HIV at the end of 2006, of which 20.8 million were found in Sub-Saharan Africa. Another 11.8 million persons have already died from the disease since the beginning of the epidemic, mostly in Africa. And about 350,000 infants now become infected each year, about 90 percent of whom are African children. Overall, more than 90 percent of the new infections each year are found in the developing countries. In many regions of the world, new HIV infections are heavily concentrated among young people (15 - 24) years of age. Among adults 15 years and older, young people accounted for 40% of new HIV infections in 2006 (UN AIDS).

Sub-Saharan Africa continues to bear the brunt of the global epidemic. Almost half (52.7%) of all adults and children with HIV globally live in Sub-Saharan Africa, with its epicenter in Southern Africa. Nearly one third (32%) of all people with HIV globally live in southern Africa and 34 % of all deaths due to AIDS in 2006 occurred there.

AIDS probably started to spread in Ethiopia in the early 1980's. The first evidence of HIV infection was found in 1984, and the first AIDS case was reported in 1986. Although HIV prevalence was very low in Ethiopia during the

early 1980's, it has been increasing rapidly in the past few years. According to the Ministry of Health (MOH) approximately 3.2 million Ethiopians are living with HIV/AIDS. UN AIDS estimated a total of 2.1 million at the end of 2001, with an adult prevalence of 6.4 percent. The cumulative number of AIDS-related deaths from the beginning of the epidemic was estimated at about 1.2 million in 2000 and is expected to increase to 1.7 million by 2005 (MOH, 2000).

Addis Abeba is, of course, the capital city of the Federal Democratic Republic of Ethiopia and the major urban center in the country. The Central Statistical Agency estimates the 2005 population of Addis Ababa at about 2.8 million persons, with a population growth rate near 2.9 percent per year. Much of the population growth in the city still stems from migration from the countryside and smaller urban areas. Unemployment is high and incomes are low. A recent report indicated that 60 percent of households earn less than 300 Birr per month. The city is also characterized by substandard housing conditions, high infant and maternal mortality rates, inadequate health services and poor sanitation. The presence of large numbers of commercial sex workers aggravates the spread of HIV and other sexually transmitted diseases.

In Addis Ababa, approximately one out of six adults is already infected with the virus, reflecting much higher infection rates than in many other parts of the country. An epidemic of this magnitude threatens the health and development of the entire country. (Source: http: // www. hivaidsclearing house. unseco. org).

Voluntary counselling and testing (VCT) service plays a significant role in the prevention of HIV/AIDS especially in the all rounded endeavor to provide care and support for people living with HIV/AIDS.

Now, a collaborative effort is being undertaken by governmental and non-governmental organizations to make the VCT service widely available to the population at large in addition to the social, economic, and psychological support it gives for people living with HIV/AIDS and AIDS orphans. The organization for

social services for AIDS (OSSA) has been actively involved in activities to increase access to VCT services to the society with the financial support provided by the President's Emergency Plan for AIDS Relief (PEPFAR) through the United States Center for Disease Control and Prevention (CDC). In an effort to make counselling and blood test service accessible to the society in Addis Ababa, a special mobile HIV counselling and testing service was inaugurated on September 9, 2005 with the financial and technical support from PEPFAR through CDC and it has been giving service since then. The mobile counselling and testing service gives reliable service to clients similar to those provided in health facilities and model VCT centers, and what makes it special is that it makes quality and fast services available where the clients are residing.

Having seen the above facts, it is worthy to see some of the impacts of demographic and HIV related risk behavior variables on the status of HIV/AIDS for the case of OSSA mobile VCT in urban Addis Ababa.

1.2 Socio-demographic profile among VCT clients

VCT counselling and testing is a cornerstone for the early access to prevention as well as care and support services. In Ethiopia, at national adult HIV prevalence of 7.3% and urban prevalence of 13%, common demand for VCT is gradually increasing (Bekele, 2000). Prevalence of HIV infection among VCT clients varied according to socio-demographic variables.

HIV infection levels increase directly with education among both women and men and are markedly higher among those who have a secondary or higher education compared with those with less education. Employment (in the past 12 months) is also related to HIV levels among both women and men, with those who are employed being more likely than the unemployed to be infected. Particularly among men, those who were unemployed during the 12-month period prior to the survey are heavily concentrated in the younger age groups where HIV levels are quite low (CSA, Demographic and Health Survey, 2005).

The Annual report on AIDS in Ethiopia for the Ethiopian Fiscal year 1998 (Hamle 1, 1997 - Sene 30, 1998 E.C) indicates that a total of 564351 VCT clients received counselling and testing services. The proportion of males and females was nearly equal. About 38.2% were reported from Addis Abeba. The overall HIV prevalence among VCT clients was 13.7% (15.7% among females and 11.6% among males). The lowest HIV prevalence was reported from Benishangul Gumuz and the highest from Somali region (AIDS in Ethiopia, sixth report, 2006).

A study on VCT clients composition among VCT clients attending the Family Guidance Association of Ethiooia (FGAE) was done for the 30257 VCT clients using the multi-level logistic regression model in the analytic sample. Approximately half of the men were never-married compared to just under 70 percent of the women. Both men and women in the sample were relatively well educated. Women were much more likely to be unemployed than men while men are more likely to be in manual skilled or professional positions. HIV infection is higher among female than male clients (Bradlrey, et al, 2007).

During the specified period of about twenty-seven months, a total of 6599 clients have received VCT services from the Kassanchis Health Center. On the average there were about 244 VCT clients per months or 12 clients per day. The distribution of VCT clients by some selected socio demographic variables shows that among the 6599 clients a little over half (54.1%) were females. With the overall average age of 27.4 years ($\pm$8.6), male clients appeared to be older by about four years than their female counterparts (29.7 years vs 25.5 years). The majority of clients were in the age range of 15-24 years (42.4%) and 25-34 years (41.1%). Among the married and never-married clients, the number of males was found to be a little higher than that of the females while among those whose marriage was dissolved (separated /divorced and widowed) the proportion of female clients is higher than males, particularly among widowed clients (79%).

Nearly half of the clients (48.8%) have attained secondary school and females appeared to be less educated than males. Unemployed clients accounted for about two-thirds of the total clients and 66.6% of the total unemployed clients were females. (Antenane Kora et al., 2005).

HIV infection rates are much higher in urban than rural areas. This pattern is consistent with the concentration of vulnerable groups and the high prevalence of the high-risk sexual behavior in towns.

1.3 OBJECTIVES OF THE STUDY

General objectives

♣ To investigate the impacts of some demographic and HIV related risk behavior variables on HIV/AIDS status of an individual.

♣ To develop a statistical model that predicts the HIV/AIDS status of an individual on the knowledge of some demographic and HIV related risk behavior variables.

Specific objectives

- To examine the relationship between the covariates and the dependent variable (HIV/AIDS status).

- To comprehend and describe the perception of HIV/AIDS and attitudes to VCT for HIV among young people in Addis Abeba.

- To show how to reduce risk behavior amongst clients testing in VCT.

- To identify the factors and variables that describe the relationship better, and

- To provide information to government and other concerned bodies in setting HIV/AIDS policies and strategies to accelerate the prevention, control, and mitigation effort in Ethiopia.

1.4 Definition of Terms

HIV/AIDS counselling is a confidential process that enables individuals to examine their knowledge and behaviors in relation to their personal risk of acquiring and transmitting HIV infection

Voluntary Counselling and Testing (VCT) is defined as client-initiated HIV testing to learn HIV status that includes both voluntary Pre-test Counselling session and Voluntary HIV Testing.

Pre-Test Counselling Session is defined as the counselling and education session before the test is done. Issues covered in a typical counselling session include a discussion of demographic profile of the client and what type of risky behavior the client might have had in the past, exploring why the client is seeking an HIV test that day, and determining if the client is emotionally stable enough to handle a positive result. All conversations taking place in a counselling session are completely confidential. Some of the questions that the counselor asks the client are: - history of STI, educational level, employment status, marital status, suspected exposure time, ever-had sex with penetration, …etc.

HIV status is the result that the client hears from the counsellor after giving a sample of blood in laboratory. In this study only positive result and negative results are included. Those clients whose result is indeterminate are excluded from the study.

Source of the above definitions: - UNAIDS May 2000 "Voluntary counselling and testing" best practice collection, technical update Geneva UN/AIDS.

1.5. Possible applications of the result

- The results help in designing appropriate HIV/AIDS policies and strategies that contribute to the prevention of the epidemic.
- The results help to understand the relationship between Pre-Test Counselling Session and HIV/AIDS status.
- The results help as a basis for further study in the area.

1.6. Limitations

In this study only positive HIV status and negative HIV status are included. Clients whose results are indeterminate are excluded from the study.

CHAPTER TWO

Literature Review

In most of the studies done before different analyses (method of estimation) were applied to examine the effects of socio-demographic characteristic and behaviors in relation to personal risk of acquiring and transmitting HIV. In this respect, an attempt was made to summarize the literature about these methods and give some image regarding the impacts of selected characteristics on status of HIV.

2.1 Voluntary counselling and testing for HIV

VCT for HIV is important in reducing sexual risk behavior (Muller et al., 1992; Painter, 2000). Weinhardt, et al. (1999) reviewed 27 published papers where sexual behavior was compared before and after testing. Among HIV positive individuals, a reduction in sexual risk-taking was reported. A similar behavioral change was not seen among HIV negative individuals. In fact, HIV negative individuals did not significantly change behavior, compared to those who were untested. The Voluntary HIV counselling and Testing Efficacy study using a stratified cross-sectional survey in Kenya, Tanzania and Trindad (Coates et al., 2000) showed similar results with the effect mainly in secondary prevention. The influence of VCT on people's behavior was investigated in rural Uganda using the same method in 1993: participants in a VCT program did not display any differences in condom use and number of casual partners after three months when compared to non-participants (Kipp et al., 2001). A more recent study from rural Uganda shows the same lack of change in risk behavior after VCT for HIV (Matovu et al., 2005).

Acceptability of VCT for HIV in developing countries has been studied in pregnant women using a cross-sectional descriptive study (Cartoux et al., 1998,

De Paoli et al., 2004, Pool et al., 2001) and in different populations of adults (Fylkesnes and Siziya, 2004, Maman et al., 2001,Van Dyk, 2003). Pregnant women generally have positive attitudes towards VCT for HIV, especially if they see clear benefits such as prevention of mother-to-child transmission (De Paoli et al ., 2004). The obstacles to VCT are fear of rejection by spouses and domestic violence as well as fear of being mistreated or even killed by maternity staff (Pool et al., 2001). For adults, important predictors of testing are costs, physical availability and the link of VCT to treatment opportunities. VCT is often out of reach for youths and women. Young people are especially vulnerable and may feel reluctant to go for HIV testing (Nuwaha et al., 2002). Fear of social rejection if the result become HIV positive and fear of being seen by others at the service center are mentioned by a significant number of respondents as reasons for not willing to be tested for HIV.

In rich countries, where anti retroviral (ARV) drugs are available, youths with an increased risk of infection are more likely to be tested for HIV (Main et al., 1994, Samet et al., 1997). Similarly, a study on HIV testing behavior among Canadians using multinomial logistic regression model indicated that those at risk are more likely to be tested (Houston et al., 1998).

Young people have a need for adequate information and appropriate counselling so that they can make informed choices about their sexuality and reproductive health. A study in Zimbabwe using a simple logit regression (Kim et al., 1997) revealed that young people are likely to ask question about sexual matters, and they have concerns about privacy and confidentiality in the counselling situations. Many providers believe that parents should be notified if young unmarried clients make young people reluctant to seek advice at healthcare facilities. There is a lack of qualitative information about young peoples' attitudes to VCT for HIV in East Africa (Amuyunzu-Myamong et al., 2005).

A study conducted from 2005-2006 in Ethiopia, with regard to the effect of HIV-VCT on HIV incidences among a cohort of participants using Non-parametric methods, univariate and multivariate logistic regression methods indicated that HIV-VCT utilization decreases the incidence of HIV infection (Sahilu T., et al 1999). Recent studies have shown that VCT is a cost-effective intervention for reducing HIV-related risk behaviors, particularly when applied to risk-couples (Allen E and Glynn E 2001). VCT services are also necessary for the prevention of mother-to-child transmission (MTCT) of HIV and to identify women and families who may benefit from MTCT interventions. In the absence of VCT services, most women in Africa have no definitive way of knowing their HIV status until they themselves fall ill with identifiable symptoms of AIDS, or until they give birth to a baby who is diagnosed with the virus and eventually dies from AIDS. For women identified as being HIV-positive before or during pregnancy, test related counselling could help women plan for their future and the future of their families.

2.2 HIV Related risk behaviors

A behavioral study conducted in 2002 in all regions of Ethiopia using Hierarchical regression analysis indicated that nearly two third of out-of school youth reported that they were sexually active and had sex with two or more partners in the last two years (Mitike G. et al., 2002). Despite high level of knowledge, a significant proportion of the population, particularly the youth are at high risk of HIV infection. Sex workers and their clients are the most vulnerable population to HIV/AIDS. However, the illicit nature of the industry in many African countries makes it an exceptionally difficult community to reach. Though some interventions such as the female condom have shown promise, HIV continues to spread rapidly among this group (Corey L et al., 2004).

Behavioral factors, such as drinking alcohol and smoking were also connected to higher HIV risk. A study on factors affecting accessibility and acceptability of

VCT service for HIV/AIDS in Bahirdar town by Family Guidance Association of Ethiopia (FGAE) in 2001 using a community-based cross sectional survey showed that the never-married and those who did not reveal their martial status, indicated that about 23% have lovers (boy/girl friend) and out of whom 73.8% said to have had experienced sex with their partner. About 25% of the respondents revealed that they had sex with other persons other than their regular partners and about 16% reported that they have more than one sexual partner at the moment (FGAE).

A study by Antenane Kora et al., (2005) at the CARE-Ethiopian VCT center at the Kassanchis Health post using multivariate analysis, mainly logistic regression indicates that the overwhelming majority of clients (82%) ever-had sexual practices in the past and female clients were less likely than their male counterparts to report sexual practices. About 69% of the total clients had over six months suspected exposure time and 5.8% had no exposure. Condom use among sexually active clients was found to be low. The majority 74.7% of the clients reported that they did not use condom in the three months time prior to testing and during last sex. High prevalence of HIV is observed among those ever-married, the less educated, the unemployed clients and also among clients who had started sex, who never used condoms and among those who had a history of STIs. Generally the findings suggest that HIV affects people of low socioeconomic status and who are exposed to unpleasant sexual practices. Exposures happening due to other factors than sexual practices have also contributed to HIV infection. This was demonstrated by the observed infection rate of 6.5% among clients with no sexual exposures (Antenane Kora, et al., 2005).

HIV transmission is largely due to heterosexual transmission. Although some groups such as sex workers are more severely affected than others, HIV affects the general population and is not just a problem that is limited to minorities or vulnerable groups. The male to female ratio is approximately 1:1 women have

higher rates of infection in the younger age groups than males (UN AIDS). HIV risk is also assumed to increase with the number of lifetime sexual partners that an individual has. This suggests that HIV risk does not rise directly with the number of sexual partners but that having a large number of partners (five or more for women and ten or more for men) is associated with significantly higher rates of HIV infection (CSA, Demographic and Health Survey, 2005).

CHAPTER THREE

Data and Methodology

3.1 The Data

The data used in this study come from the special records kept at the Biruh-Hiwot (OSSA) mobile counselling and testing service center in Addis Abeba. The mobile voluntary counselling and testing data have been recorded since the inception of the service in September 9, 2005 using a standard recording format developed by the US Center for Disease Control and Prevention (CDC). The recording formats were filled by the VCT center counsellors at the time of pre and post-test counselling and entered into computer at the project office. The study subjects are men and women clients who visited the center for VCT service between September 2005 and April 2007. A total of 14,810 people received the service during the specified period and records of all these clients were used for this study. The data are analyzed using SPSS software Packages.

3.2 Variables of interest

1. The Dependent variable

The response /dependent variable in this study is HIV status of an individual which is dichotomized as 1 if he /she is positive and as 0 if he /she is negative or HIV status of an individual (0=Negative, 1=Positive).

2. The Independent Variables

The independent variables that are used in this study are classified as demographic variables and HIV related risk behavior variables. Variables such as age, sex, marital status, etc... are considered as demographic variables, while "reason why client here today", previously tested, ever had sex with penetration etc... are considered as HIV related risk behavior variables.

Table 3.1: Independent variables included in the analyses

Variable designation	Description	Value labels
	Demographic Variables	
X_1	Age in year	0= < 15,1=15-30, 2=31-49, 3= ≥ 50
X_2	Sex	0= Male, 1= Female
X_3	Marital status	0=Never-married, 1= married, 2= Divorced, 3 = Widowed
X_4	Educational level	0= Illiterate 1= Grades (1-6) 2= Junior and High school, 3= ≥12 (tertiary)
X_5	Employment status	0=Inactive, 1= Active
X_6	Occupation	0=Student, 1=Elementary occupation / unskilled 2=Professional, 3= Merchant, 4= Others
	HIV related risk behavior variables	
X_7	The most important reason why client is here today	0=Client risky, 1=Visa applicant, 2=Plan for future, 3=Other
X_8	Previously tested	0=No, 1=Yes
X_9	Ever had sex with penetration	0=No, 1=Yes
X_{10}	Suspected exposure time	0= < one month, 1=1 to 3 months, 2= 4 to 6 months 3= over 6 months, 4= not exposed
X_{11}	Condom use last 3 months	0=Never, 1=Always, 2=Sometimes, 4=Not applicable*
X_{12}	Used condom last time had sex	0=No, 1=Yes, 2=Doesn't remember, 4=Not applicable*
X_{13}	History of STI	0= Yes, 1= other (No, Don't Know)

*Refers to those clients who did not have sex with penetration

The method for specifying the design variables involves setting all of them equal to zero for the reference group, and then setting a single design variable equal to 1 for each of the other groups. This method, sometimes referred to as " reference cell coding," and is used to form a set of design variables to represent the categories of the variables. The SPSS output for categorical variable coding is shown in Table 3.2

Table 3.2 The SPSS coding for Categorical Variables

		Frequency	Parameter coding			
			(1)	(2)	(3)	(4)
Suspected exposure time	Less than one month	1104	1.000	.000	.000	.000
	1 to 3 months	1124	.000	1.000	.000	.000
	4 to 6 months	1352	.000	.000	1.000	.000
	Over 6 months	6056	.000	.000	.000	1.000
	Not exposed	5078	.000	.000	.000	.000
Occupation	Student	2847	1.000	.000	.000	.000
	Elementary occupation / unskilled	3058	.000	1.000	.000	.000
	Professional	1529	.000	.000	1.000	.000
	Merchant	1329	.000	.000	.000	1.000
	Other	5951	.000	.000	.000	.000
The most important reason why client is here today	Client risky	7272	1.000	.000	.000	
	Visa applicant	35	.000	1.000	.000	
	Plan for future	6585	.000	.000	1.000	
	Other	822	.000	.000	.000	
Marital status	Never-married	9117	1.000	.000	.000	
	Married	4374	.000	1.000	.000	
	Divorced	944	.000	.000	1.000	
	Widowed	279	.000	.000	.000	
Educational level	Illiterate	724	1.000	.000	.000	
	Grade (1 - 6)	3179	.000	1.000	.000	
	Junior and high School	8441	.000	.000	1.000	
	> 12 (Tertiary)	2370	.000	.000	.000	
Age in year	Less than 15 years	16	1.000	.000	.000	
	15 - 30 years	9969	.000	1.000	.000	
	31 - 49 years	3871	.000	.000	1.000	
	50 years and above	858	.000	.000	.000	
Condom use last 3 months	Never	7679	1.000	.000	.000	
	Always	2258	.000	1.000	.000	
	Sometimes	1162	.000	.000	1.000	
	Not applicable	3615	.000	.000	.000	
Used condom last time had sex	No	8629	1.000	.000	.000	
	Yes	3063	.000	1.000	.000	
	Does not remember	30	.000	.000	1.000	
	Not applicable	2992	.000	.000	.000	
Sex	Male	9956	1.000			
	Female	4758	.000			
Previously tested	No	9728	1.000			
	Yes	4986	.000			
Employment status	Inactive	4745	1.000			
	Active	9969	.000			
Ever had sex with penetration	No	3218	1.000			
	Yes	11496	.000			
History of STI	Yes	1054	1.000			
	Other (No, don't know)	13660	.000			

3.3 Methodology

3.3.1 The Logistic Regression

Binomial (or binary) logistic regression is a form of regression, which is used when the dependent variable is dichotomous, such as presence/absence or success/failure, HIV positive/HIV negative and the independent variables are of any type. Multinomial logistic regression exists to handle the case of dependent variables with more than two classes. When multiple classes of the dependent variable can be ranked, then ordinal logistic regression is preferred to multinomial logistic regression. Logistic regression can be used to predict a dependent variable on the basis of continuous and/or categorical independent variables and to determine the percent of variance in the dependent variable explained by the independent variables; to rank the relative importance of the independent variables; to assess interaction effects; and to understand the impact of the covariate control variables.

3.3.2 The model

The basic aim of modeling is to derive a mathematical representation of the relationship between an observed response variable and a number of explanatory variables, together with a measure of the inherent uncertainty of any such relationship. Statistical models constructed for response variables are at best an approximation to the manner in which an observable variable depends on other variables. No statistical model can be claimed to represent truth and, by the same token, no one model can be termed the correct model. Some models will be more appropriate than others, but typically, for any set of data, there will be a number of models, which are equally well suited to the purpose in hand, and the basis for choosing a single model from amongst them will not rest on statistical grounds alone. Statistical models are essentially descriptive and, in as much as they are based on experimental or observational data, may be described as empirical data. (Collett 1991)

The dependent variable in logistic regression is usually dichotomous, that is, the dependent variable can take the value 1 with a probability of success θ, or the value 0 with a probability of failure 1- θ. This type of variable is called a Bernoulli (or binary) variable.

The independent or predictor variables in logistic regression could be discrete, continuous or a mix of both. Logistic regression makes no assumption about the distribution of the independent variables. They do not have to be normally distributed, linearly related or have equal variance within each group. In logistic regression the relationship between the predictor and response variable is not linear. Instead, the logistic regression function is used, which is the logit transformation of θ. In logistic regression, a single outcome variable Y_i (i=1,...,n) follows a Bernoulii probability function that takes on the value 1 with probability θ_i and 0 with probability 1- θ_i. Then θ_i varies over the observations as an inverse logistic function of a vector X_i, which includes a constant and *k* explanatory variables, that is,

$$\theta(x) = \frac{e^{(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki})}}{1 + e^{(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki})}}$$

where β_0 is the constant of the equation and, β_i's are the coefficients of the predictor variables.

An alternative form of the logistic regression equation is:

$$\text{logit}[\theta(x)] = \log\left[\frac{\theta(x)}{1-\theta(x)}\right] = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} = x^{/}\beta$$

The coefficients can be interpreted as the change in the log-odds associated with a one unit change in the corresponding independent variable or the odd increases multiplicatively by e^{β} for every one unit change increase in x.

The ratio of success to failure, $\frac{\theta(x)}{(1-\theta(x))}$, is called the odds of success.

Through the relation $X^{/}\beta = \log\left(\frac{\theta(x)}{1-\theta(x)}\right) = \operatorname{logit}(\theta(x))$, the dependence of the success probability on the explanatory variables, the probability scale is transformed from the range $(0,1)$ to $(-\infty,\infty)$. This transformation ensures that the fitted success probabilities will lie between 0 and 1.

The goal of logistic regression is to correctly predict the category of outcomes for individual cases using the most parsimonious model. To accomplish this goal, a model is constructed that includes all predictor variables that are useful in predicting the response variable. Several different options are available during model construction. Variables can be entered into the model in the order specified by the researcher or logistic regression can test the fit of the model after each coefficient is added or deleted, called stepwise regression.

Stepwise regression is used in the exploratory phase of research but is not recommended for theory testing (Menard, 1995). Theory testing is the testing of a priori theories or hypothesis of the relationships between variables. Exploratory testing makes no a-priori assumptions regarding the relationships between the variables, thus the goal is to discover relationships. Backward stepwise regression appears to be the preferred method of exploratory analyses, where the analysis begins with a full or saturated model and variables are eliminated from the model in an iterative process. The fit of the model is tested after the elimination of each variable to ensure that the model still adequately fits the data. When no more variables can be eliminated from the model, the analysis has been completed.

There are two main uses of logistic regression. The first is the prediction of group membership. Since logistic regression calculates the probability or success

over the probability of failure, the results of the analysis are in the form of an odd ratio. In order to fit the model to the data, the parameters of the model have to be estimated. In fact, the model fitting process could be facilitated by the widely available Statistical software such as SAS, SPSS, and GLIM. In this study, SPSS is used for analyzing the data.

3.3.3 Parameter Estimation

Since the logistic regression model is nonlinear an iterative algorithm is necessary for parameter estimation. Let p be the probability of success and it is equivalent to the probability that the response variable assumes value one. Then

$$P(Y=1)=\frac{1}{1+e^{-x'\beta}}$$

each observation (response) can be considered as an outcome of a Bernoulli trial. Hence for the ith observation Y_i the Bernoulli distribution is $P(Y=y_i)=p^{y_i}(1-p)^{1-y_i}$.

Then the likelihood function is the joint probability distribution of all *n* observation is

$$L=\prod_{i=1}^{n} p^{y_i}(1-p)^{1-y_i}=\prod_{i=1}^{n}\left(\frac{1}{1+e^{-x'\beta}}\right)^{y_i}\left(\frac{e^{-x'\beta}}{1+e^{-x'\beta}}\right)^{1-y_i}$$

Taking the natural logarithm of both sides yields

$$\ln L=\sum_{i=1}^{n} y_i\ln\left(\frac{1}{1+e^{-x'\beta}}\right)+\sum_{i=1}^{n}(1-y_i)\ln\left(\frac{e^{-x'\beta}}{1+e^{-x'\beta}}\right)$$

Hence, through maximization of the above log-likelihood function we can theoretically estimate the parameter vector β. But the equation is nonlinear in β, and as a result the estimates do not have a closed form expression. Therefore, β will be obtained by maximizing using a numerical iterative method (Agresti, 1996).

3.3.4 Assessing the goodness of fit

Measures of goodness of fit are stastical tools used to explore the extent to which the fitted response obtained from the postulated model compares with the observed data. Clearly, the fit is good if there is a good agreement between the fitted and the observed data. In practice, any model may not exactly represent the true relationship between θ(x) and x. Thus, as the sample size increases, the model-based estimator may not converge exactly to the true value of the probability. As explained above logistic regression model was used to describe and make inferences about the effects of predictors on the response. There is no guarantee, however, that a particular model of this form is appropriate or that it provides a good fit to the data. So, we need to assess the goodness of fit of the model (Agresti, 1996). The Pearson's Chi-square, the likelihood ratio test (LRT)(Deviance), Hosmer- Lemeshow test and the Wald test are the most commonly used measures of goodness of fit for categorical data. The following section will give a brief discussion on how each of this criterion can be used as a measure of goodness of fit of a logistic model.

3.3.4.1 The likelihood ratio criterion (Deviance analysis)

A "likelihood" is a probability, specifically the probability that the observed values of the dependent variable may be predicted from the observed values of the independent variables. Like any probability, the likelihood varies from 0 to 1. The log likelihood (LL) is its log and varies from 0 to infinity (it is negative because the log of any number less than 1 is negative). LL is calculated through

iteration, using maximum likelihood estimation (MLE). Log likelihood is the basis for tests of a logistic model. The deviance D is defined as

$$D=-2\ln\left[\frac{likelihood of the current \bmod el}{likelihood of the saturated \bmod el}\right]=-2\left[\frac{LL_{current}}{LL_{saturated}}\right]$$

When one model contains terms that are additional to those in another, the two models are said to be nested. Deviance can be used to compare two nested models for grouped as well as ungrouped binary data. The statistic D therefore measures the extent to which the current model deviates from the full model and is termed the deviance. This statistic is a useful measure for detecting lack of fit of a given logistic regression model for grouped binary data. However, it proves to be uninformative for assessing lack of fit of a model for individual binary data (Collet, 1991). The difference in the deviances of two nested models measure the extent to which the additional terms improve the fit of the model to the observed response variable. If we compare the two nested models:

Model (1): logit (p)= $\beta_o+\beta_1X_1+\beta_2X_2+...+\beta_hX_h$

Model (2): logit (p)= $\beta_o+\beta_1X_1+\beta_2X_2+...+\beta_hX_h+\beta_{h+1}X_{h+1}+...+\beta_kX_k$

Let the deviance under each model be D_1 with degree of freedom $\mu_1=(n-h-1)$ and D_2 with degree of freedom $\mu_2=(n-k-1)$ respectively, where n is the number of binomial observasions. Model (1) is nested within Model (2) since Model (2) contains more terms than Model (1). The difference in deviance D_1-D_2 will reflect the combined effect of the variables $X_{h+1},...,X_k$ after $X_1, X_2,..., X_h$ have already been included in the model. This difference is described as the deviance of fitting $X_{h+1},...,X_k$ adjusted for elliminating $X_1, X_2,..., X_h$. Since the deviance for each model has an approximate χ^2 distribution, the difference between two deviances can also be approximated by χ^2 distribution. D_1 has a χ^2 distribution with degree of freedom μ_1, D_2 has χ^2 distribution with degree of freedom μ_2 and D_1-D_2 has χ^2 distribution with degree of freedom $\mu_1-\mu_2$. This is based on the likelihood function of the observed $\hat{p}_i$ for the fitted model ($\hat{L}_c$), and the likelihood

function for the true success probability under the assumption of full model like ($\hat{L}_f$). The deviance denoted by D_1 and have the distributions:

$$D_1 = -2\left[\log \hat{L}_{c_1} - \log \hat{L}_f\right] \sim \chi^2(\mu_1) \text{ and}$$

$$D_2 = -2\left[\log \hat{L}_{c_2} - \log \hat{L}_f\right] \sim \chi^2(\mu_2)$$

However, the deviance for individual binary data cannot be approximated by χ^2 distribution. This is because of the inclusion of the likelihood ratio under the full model. But when comparing two deviances, the term involving $\hat{L}_f$ disappears

and $$D_1 - D_2 = -2\left[\log \hat{L}_{c_1} - \log \hat{L}_{c_2}\right] \sim \chi^2(\mu_1 - \mu_2)$$

This is the usual likelihood ratio test for comparing two models.

Large values are encountered when $\hat{L}_c$ is small relative to $\hat{L}_f$, indicating that the current model is poor. On the other hand, small values of D are obtained when $\hat{L}_c$ is similar to $\hat{L}_f$, indicating that the current model is good one.

3.3.4.2 Pearson's X^2 - statistic

This statistic is defined by

$$X^2 = \sum_{i=1}^{n} \frac{\left(Y_i - n_i \hat{p}_i\right)^2}{n_i \hat{p}_i \left(1 - \hat{p}_i\right)}$$

where $$\hat{p}_i = \frac{e^{\hat{\eta}_i}}{1 + e^{\hat{\eta}_i}}, \qquad \hat{\eta}_i = \sum_{j=1}^{k} \hat{\beta}_j X_{ji}$$

The linear predictor, is obtained by substituting the Maximum Likelihood Estimation (MLE) of the β_j in η_j. The distribution of X^2 is asymptotically Chi-square with $(n - k - 1)$ degrees of freedom. Large values of X^2 can be taken as evidence that the model does not adequately fit the data. Because the model parameters are estimated by the method of maximum likelihood, it is recommended that one uses the Deviance statistic as a criterion for goodness of fit of the logistic regression model.

3.3.4.3 Hosmer-Lemeshow Test

The test divides subjects into deciles based on predicted probabilities, and then computes a chi-square from observed and expected frequencies. Then a probability (p) value is computed from the chi-square distribution with 8 degrees of freedom to test the fit of the logistic model. The Hosmer-Lemeshow test statistics is given by

$$\hat{C} = \sum_{i=1}^{10} \frac{(O_k - E_k)^2}{V_k}$$

where O_k is the observed number of events in the kth group, E_k is expected number of events in the kth group, and V_k is a variance correction factor for the kth group. If the H-L goodness-of-fit test statistic is greater than 0.05, we will not reject the null hypothesis that there is no difference between observed and model-predicted values, implying that the model estimates fit the data at an acceptable level. That is, well fitting models show nonsignificance on the H - L goodness-of-fit test, indicating that model-prediction is not significantly different from the observed values. This does not mean that the model necessarily explains much of the variance in the dependent, only that however much or little it does explain is significant. As the sample size gets large, the H-L statistic can find smaller and smaller differences between observed and model-predicted values to be significant.

3.3.4.4 Test of goodness of Fit of the Final Model - Wald Test

A Wald test is used to test the statistical significance of each coefficient (β) in the model. The test statistic is a chi-square statistic with a desirable outcome of non-significance, indicating that the model prediction does not significantly differ from the observed.

The hypothesis to be tested is:

$H_0 : \beta_j = 0$ versus $H_A : \beta_j \neq 0$ at α level of significance.

The Wald test statistics, Z, for this hypothesis is

$$Z^2 = \frac{\hat{\beta}_j^2}{Var\left(\hat{\beta}_j\right)} \sim \chi^2(1)$$

$\hat{\beta}_j^2$ is the square of the estimated regression coefficient and Var$\left(\hat{\beta}_j\right)$ is the variance of $\hat{\beta}_j$. However, several authors have identified problems with the use of Wald statistic. Menard (1995) warns that for large coefficients, the standard error is inflated, lowering the Wald statistic (chi-square) value. The likelihood-ratio test is more reliable for small sample sizes than the Wald test (Agresti, 1996).

Assumptions

Logistic regression is popular because it enables the researcher to overcome many of the restrictive assumptions of OLS regression:

1. Logistic regression does not assume a linear relationship between the dependent and the independent variables. It can handle nonlinear effects even when exponential and polynomial terms are not explicitly added as additional independent variables because the logit link function on the left-hand side of the logistic regression equation is non-linear. However, it also possible and permitted to add explicit interaction and power terms as

variables on the right-hand side of the logistic equation, as in OLS regression.

2. The dependent variable need not be normally distributed (but does assume its distribution is within the range of the exponential family of distributions, such as normal, Poisson, binomial, gamma).
3. The dependent variable need not be homoscedastic for each level of the independent variables; that is, there is no homogeneity of variance assumption: variances need not be the same within categories.
4. Normally distributed error terms are not assumed.
5. Logistic regression does not require that the independent variables be interval scaled.
6. Logistic regression does not require that the independent variables be unbounded.

However, other assumptions that still apply are:

1. Meaningful coding. Logistic coefficients will be difficult to interpret if not coded meaningfully. The convention for binomial logistic regression is to code the dependent class of greatest interest as 1 and the other class as 0.
2. Inclusion of all relevant variables in the regression model: If relevant variables are omitted, the common variance they share with included variables may be wrongly attributed to those variables, or the error term may be inflated.
3. Exclusion of all irrelevant variables: If causally irrelevant variables are included in the model, the common variance they share with included variables may be wrongly attributed to the irrelevant variables. The more the correlation of the irrelevant variable(s) with other independent variables, the greater the standard errors of the regression coefficients for these independent variables (Source:- www.http // file:F:/PA 765 Logistic Regression.html)

CHAPTR FOUR

DATA ANALYSIS

The data are analyzed using the Statistical Package for Social Science (SPSS) version 13 and STATA 9.2. Summary statistics are used to describe a set of observations, the Pearson Chi-square test was conducted to assess the association between variables and logistic regression analysis was run to assess the differences among the predictor categories and dependent variable.

4.1 Summary Statistics

A common first step in data analysis is to summarize information about variables in your data set. Descriptive statistics are used to summarize statistical features of a set of observation in order to communicate as much as possible as simple as possible.

4.1.1 Demographic profile:

During the specified period of about 19 months, a total of 14,810 clients received VCT services from the OSSA mobile VCT center in Addis Abeba. On average there were about 26 clients per day. Table 4.1 presents the distribution of VCT clients by some demographic variables. Among the 14,810 clients a majority (67.6%) were males. With the overall average age of 28.8 years ($\pm$10.5), male clients appeared to be older by about three years than their female counterparts (29.7 years vs. 26.9). The majority of clients were in the age range of 15-30 years (67.7%) and 31-49 years (26.3%). About (62 %) of the total clients were never-married followed by currently married clients (29.7 %). Among the married and never-married clients, the number of males was found to be higher than that of the females while among those whose marriages was dissolved (divorced and

widowed) the proportion of female clients is higher than males, particularly among widowed clients (69.6%). About (57.1 %) of the clients have attended junior and high school and females appeared to be less educated than males. Active clients accounted for two-thirds (67.7%) of the total clients and 74.4% of the total active clients were males. The occupations of the majority of the clients were others (housewife, armed forces...etc) followed by those who have elementary occupation/unskilled.

Table 4.1 **Percentage distributions of VCT clients by some demographic variables**

Demographic variables		Number	Percentage
Sex	Male	10024	67.7
	Female	4786	32.3
Age in year	< 15	16	0.1
	15-30	10028	67.7
	31-49	3899	26.3
	≥ 50	863	5.8
Marital status	Never-married	9177	62.0
	Married	4396	29.7
	Divorced	951	6.4
	Widowed	280	1.9
Education	Illiterate	726	4.9
	Grade (1-6)	3185	21.5
	Junior and high school	8460	57.1
	> 12 (Tertiary)	2377	16.0
Employment status	Inactive	4773	32.2
	Active	10025	67.7
Occupation	Student	2858	19.3
	Elementary occupation	3066	20.7
	Professional	1559	10.5
	Merchant	1341	9.1
	Other (housewife, etc)	5986	40.4

4.1.2 HIV related risky behavior

The reason cited for seeking VCT services were various, but the majority (49.4%) said they practiced risky behavior, followed by those who want to "plan for future" (44.7%), 0.2% did for visa application and the others account about 5.6% e.g. because they have "seen" symptoms of HIV/AIDS...etc. (Table4.2)

Table 4.2: **Percentage distribution of VCT clients by sexual experience, estimated exposure time, condom use and STI history by sex**

Category		Male % (n)	Female % (n)	Total % (n)
The most important reason why client here today	Client risky	51.3 (5142)	45.3 (2169)	49.4(7311)
	Visa applicant	0.1 (13)	0.5 (24)	0.2(37)
	Plan for future	43.7(4385)	33.8(2242)	44.7(6627)
	Other	4.8 (484)	2.4(351)	5.6(835)
Previously tested	No	66.9 (6702)	64.2 (3072)	66(9774)
	Yes	33.1 (3322)	35.8 (1714)	34(5036)
Ever-had-sex with penetration	No	19.3(1929)	27.3(1305)	21.8(3234)
	Yes	80.7(8091)	72.7(3480)	78.2(11571)
Suspected exposure time	< 1 months	7.3 (727)	8.1 (388)	7.5 (1115)
	1 to 3 months	8.5 (856)	5.9 (280)	7.7 (1136)
	4 to 6 months	10.2 (1022)	7.1 (339)	9.2 (1361)
	Over 6 months	41.9 (4197)	39.5 (1890)	41.1 (6087)
	Not exposed	32.1 (3222)	39.5 (1889)	34.5 (5111)
Condom use last 3 months	Never	49.0 (4914)	58.7 (2810)	52.2 (7724)
	Always	18.9 (1899)	7.9 (377)	15.4 (2276)
	Sometimes	9.6 (958)	4.4 (212)	7.9 (1170)
	Not applicable	22.5 (2252)	29 (1385)	24.6 (3637)
Used condom last time had sex	No	56.2 (5629)	63.8 (3053)	58.6 (8682)
	Yes	25.6 (2567)	10.8 (517)	20.8 (3084)
	Does not remember	0.2 (24)	0.1 (7)	0.2 (31)
	Not applicable	18 (1802)	25.2 (1207)	20.3 (3009)
History of STI	Yes	8.7 (871)	3.9 (189)	7.2 (1060)
	Other (No, Don't know)	91.3 (9153)	96.1 (4597)	92.8 (13750)

The vast majority of clients (78.2%) ever-had-sexual practices in the past and female clients were less likely than their male counterparts to report sexual

practices (27.3% vs. 19.3%). About 41.1% of the total clients had over six months suspected exposure time (time they are exposed to unsafe sex practices) and 34.5% had no exposure. Condom use among sexually active mobile VCT clients was found to be low. Nearly half (52.2%) of the clients reported that they did not use condoms in three months time prior to testing and during last sex (58.6%). Females are more likely than males for not using condoms (58.7% vs.49.0%; last three months and 63.8% vs. 56.2% last time had sex) (p=0.000). About 7.2% of the clients had had a history of STIs and males are more likely than females in reporting history of STI (8.7% vs. 3.9%) (p=0.000). Almost two third (66%) of the clients were not tested previously and male clients are more likely than female clients in not tested previously (66.9% vs. 64.2%) (p =0.000).

According to mobile VCT records of a total of 14,810 clients had their blood tested following pre-test counselling, of which about 6.4% had test results that turned out to be HIV positive. Prevalence of HIV infection varies by the socio-demographic characteristic of the VCT clients. Table 4.3 presents the distribution of sero-positive clients by some demographic/HIV related risk behavior variables. Sero-positivity is high among females than males. According to the findings, 4.8% of the male and 9.7 of the female clients were found to be HIV positive. HIV prevalence appeared to be significantly lower (5.1%) among clients 15-30 years than in clients at older age group. Clients aged (31-49) years exhibited higher prevalence of HIV infection (9.8%) and also those aged above 50 years (5.8%). HIV prevalence of VCT clients also varies by marital status. Ever-married clients are more likely to be HIV+ than never-married ones (p=0.000). HIV prevalence for never-married clients was estimated at 4.2% as opposed to married 7%.

Table 4.3: **Percentage distribution of sero-positive clients clarified by variables**

Variables	No	%
Total	**945**	**6.4**
Age in year		
<15	1	6.3
15-30	511	5.1
31-49	382	9.8
>50	50	5.8
Sex		
Male	483	4.8
Female	462	9.7
Marital status		
Never-married	385	4.2
Married	307	7.0
Divorced	167	17.6
Widowed	85	30.4
Educational level		
Illiterate	98	13.5
Grade (1-6)	268	8.4
Junior and high school	514	6.1
>12(Tertiary)	63	2.7
Employment status		
Inactive	295	6.2
Active	649	6.5
Occupation		
Student	40	1.4
Elementary occupation/unskilled	241	7.9
Professional	58	3.9
Merchant	84	6.3
Other	522	8.7
The most important reason client here today		
Client risky	655	9.0
Visa applicant	2	5.4
Plan for future	215	3.2
Other	73	8.8
Previously tested		
No	712	7.3
Yes	233	4.6
Ever had sex with penetration		
No	42	1.3
Yes	902	7.8
Suspected exposure time		
<1 month	84	7.5
1 to 3 months	68	6.0
4 to 6 months	58	4.3
Over 6 months	631	10.4
Not exposed	104	2
Condom use last 3 months		
Never	652	8.4
Always	108	4.8
Sometimes	90	7.7
Not applicable	94	2.6

Used condom last time had sex		
No	734	8.5
Yes	170	5.5
Does not remember	2	6.7
Not applicable	38	1.3
History of STI		
Yes	149	14.1
Other(No, Don't know)	796	5.8

A significantly higher prevalence was found among clients whose marriages doom ended by divorce/separation (17.6%) and death of partner or widowed (30.4%). The data further revealed decreasing pattern of HIV prevalence as client's education level increases. For instance among illiterate and Grade (1-6) clients, HIV prevalence found to be 13.5% and 8.4% respectively, whereas among junior and high school and those with tertiary educational level clients, HIV prevalence found to be 6.1% and 2.7% respectively. Likewise, inactive and active clients almost have equal chances to be HIV+ (6.2 vs. 6.5).

Levels of HIV infection are high among clients who ever-had sexual experience, who never used condom in the last three months and among those clients who do not use condom last time had sex and among those who had history of STI. Clients who ever-had sex experience have more chance to be HIV+ than with no experience (7.8% vs. 1.3%). HIV prevalence appeared to be high among clients who never used condom in the three months prior to testing (8.4%) than those who used sometimes (7.7%) and used always (4.8%) and the difference were statistically significant ($p<0.000$). Moreover, prevalence of HIV infection is significantly higher (14.1%) among clients with history of STI than others (5.8%).

4.2 Univariate Findings

Model-building strategies begin with a careful univarite analysis of each variable. The likelihood ratio chi-square test with (k-1) degrees of freedom, where k is the levels of the independent variable is exactly equal to the value of likelihood ratio test for the significance of the coefficients for the k-1 design

variables in a univariate logistic regression model that contains that single independent variable. Since the Pearson chi-square test is asymptotically equivalent to the likelihood ratio chi-square test, it is used in the analysis. The preliminary analysis was done to know which of the explanatory variables appears to have a strong association with the dependent variable. For each one of the independent variables a test of association was carried out using the Pearson Chi-Square at 5% level of significance. The software used to analyze the data is STATA 9.2 package. High values of Pearson chi-square for a given independent variable indicate that there is a strong association between each of the given independent variables and the dependent variable keeping the effect of the other factors constant. That is, testing the hypothesis:

H_o= There is no association between the dependent and an independent variable.

H_1= There is association between the dependent and a particular independent variables.

The decision was based on the Pearson chi-square value, P-value and the 5% level of significance. The bivariate association between HIV status of an individual and independent variables is shown in Table 4.4.

Table 4.4: **Bivariate association between HIV status of an individual and independent variables**

Variable label	Description	Pearson chi-square	D .f	(P-value)
X_1	Age in year	104.65	3	0.000
X_2	Sex	126.47	1	0.000
X_3	Marital status	544.38	3	0.000
X_4	Educational level	139.70	3	0.000
X_5	Employment status	0.4563	1	0.499
X_6	Occupation	203.05	4	0.000
X_7	The most important reason why	273.81	4	0.000

	client is here today			
X_8	Previously tested	39.35	1	0.000
X_9	Ever had sex with penetration	179.02	1	0.000
X_{10}	Suspected exposure time	336.89	4	0.000
X_{11}	Condom use last three months	156.52	3	0.000
X_{12}	Used condom last time had sex	198.62	3	0.000
X_{13}	History of STI	227.66	2	0.000

The above univariate findings show that HIV status of an individual is strongly associated with age, sex, marital status, educational level, occupation, the most important reason why client is at the VCT center today, previously tested, ever-had sex with penetration, suspected exposure time, condom use last three months, used condom last time, had sex and history of STI. The only independent variable that has no significance association with HIV status of an individual is employment status. Upon completion of the univariate analysis we select variables for the multivariate analysis. Any variable whose univariate test has a p-value <0.005 is considered as a candidate for the multivariate model. Accordingly, the variables that are candidates for the multivariate analysis are: - age, sex, marital status, educational level, occupation, the most important reason why client is at the VCT center today, previously tested, ever-had sex with penetration, suspected exposure time, condom use last 3 months, used condom last time had sex and history of STI. The univariate analysis demonstrates that the prevalence of HIV has association to some demographic and risk behaviors variables, and most of these variables are associated with each other. Thus, this analysis doesn't show independent effect of the variables. As a result, the dichotomous relationship between HIV prevalence and various predictor variables was analyzed using the multivariate logistic regression method.

4.3 Multivariate Findings

The main problem with any univariate approach is that it ignores the possibility that a collection of variables, each of which is weakly associated with the outcome, can become an important predictor of the outcome when taken together (Hosmer and Lemeshow, 1989). Hence, multivriate logistic regression approach that takes into account the drawback mentioned by the univariate technique is considered in this analysis. The value of the outcome variable is " The probability of contracting HIV/AIDS", which is binary discrete and is represented by "Y". The value label of the variable is "1" if the client is HIV positive and "0" if the client is HIV negative. Therefore, the following equation (after log transformation is made) is fitted.

$$Logit\,(Y) = \beta_0 + \beta_1 X_{i1} + ... + \beta_{13} X_{i3} \quad , \ i=1,2....,n$$

where $\beta_0, \beta_1, ..., \beta_{13}$ are parameters to be estimated using the maximum likelihood method in the logistic regression by defining the likelihood function.

$X_{i1}, X_{i2}, ..., X_{i13}$ are some demographic and risk behavior covariates and n represents the number of observation which is equal to n=14,810.

4.3.1 Estimates for logistic model

Fitting a model to a set of data first entails estimating the unknown parameters in the model. In order to fit a linear logistic model to a given set of data, the k + 1 unknown parameters $\beta_0, \beta_1, ..., \beta_k$ have to be estimated. These parameters are readily estimated using the method of maximum likelihood. The logistic regression using the "FORWARD STEPWISE (LIKELIHOOD)" was run. This procedure revealed that the variables " age in year ", " sex ", " marital status ", " educational level ", " employment status ", " occupation ", " the most important reason why client is here today ", " previously tested ", " suspected exposure time ", " history of STI " are important predictors of HIV status of an individual. Results obtained by the procedures of FORWARD STEPWISE (LIKELIHOOD

RATIO) are shown in Table 4.5 below. The final (optimal) logistic regression model includes only the above variables.

Table 4.5: **Estimates for the final logistic regression model**

Covariates	Sub-groups	$\hat{\beta}$	S.E$\left(\hat{\beta}\right)$	Wald	D.F	Sig	Exp$\left(\hat{\beta}\right)$
Age in year				27.444	3	0.000	
	< 15	0.691	1.077	0.412	1	0.521	1.997
	15-30	0.598	0.173	11.905	1	0.010	1.819
	31-49	0.816	0.165	24.504	1	0.000	2.262
	$\geq$ 50 (Ref)	_	_	_	_	_	_
Sex	Male	-0.690	0.079	75.633	1	0.000	0.502
	Female (Ref)	_	_	_	_	_	_
Marital status				93.340	3	0.000	
	Never married	-1.360	0.165	67.793	1	0.000	0.257
	Married	-1.293	0.156	68.594	1	0.000	0.275
	Divorced	-0.683	0.166	16.972	1	0.000	0.505
	Widowed (Ref)	_	_	_	_	_	_
Educational level				10.654	3	0.014	
	Illiterate	0.636	0.200	10.131	1	0.001	1.889
	Grade (1-6)	0.426	0.172	6.117	1	0.013	1.531
	Junior and high school	0.444	0.158	7.872	1	0.005	1.559
	> 12 (tertiary) (Ref)	_	_	_	_	_	_
Occupation				48.936	4	0.000	
	Student	-1.177	0.180	42.613	1	0.000	0.308
	Elementary-occupation	-0.271	0.088	9.497	1	0.002	0.763
	Professional	-0.445	0.165	7.257	1	0.007	0.641
	Merchant	-0.215	0.129	2.792	1	0.095	0.806
	Other (Ref)	_	_	_	_	_	_

The most important reason why client is here today				78.432	3	0.000	
	Client risky	-0.333	0.141	5.552	1	0.018	0.717
	Visa applicant	-0.556	0.761	0.533	1	0.465	0.574
	Plan for future	-1.080	0.154	49.352	1	0.000	0.340
	Other (Ref)	_	_	_	_	_	_
Previously tested	No	0.500	0.085	34.225	1	0.000	1.648
	Yes (Ref)	_	_	_	_	_	_
Ever-had sex with penetration	No	-0.676	0.202	11.144	1	0.001	0.509
	Yes (Ref)	_	_	_	_	_	_
Suspected exposure Time				49.193	4	0.000	
	Less One month	0.701	0.173	16.421	1	0.000	2.015
	1 to 3 months	0.470	0.189	6.165	1	0.013	1.600
	4 to 6 months	-0.066	0.196	0.115	1	0.735	0.936
	Over 6 months	0.687	0.143	23.001	1	0.000	1.988
	Not exposed (Ref)	_	_	_	_	_	_
History of STI	Yes	0.702	0.106	43.746	1	0.000	2.018
	(Other i.e No, Don't know) (Ref)	_	_	_	_	_	_
Constant	-3.184	-2.105	0.312	45.567	1	0.000	0.122

* Ref indicates the reference category

4.3.1.1 Variables in the Equation

Variable(s) entered in step 1: Marital status, variable entered in step 2: Suspected exposure time, variable entered in step 3: Occupation, variable entered in step 4: sex, variable entered in step 5: The most impotent reason why client is here toady, variable entered in step 6: History of STI, variable entered in step 7: Previously tested, variables entered in step 8: age, variable entered in step 9: ever-had sex with penetration, variable entered in step 10: Educational level.

The likelihood ratio test criterion is used to select and remove variables at each step of the procedures. The 0.05 level of significance is used for entry of

variables. The reference category is the last category of all independent variables.

Table 4.5 above contain the estimated coefficients ($\hat{\beta}$). The standard errors of the estimates S.E ($\hat{\beta}$) which helps in computing the values of Wald statistics W,

$$W = \left(\frac{\hat{\beta}}{S.E\left(\hat{\beta}\right)} \right)^2$$

which is given in the column " Wald " and it has a chi-square distribution with one degree of freedom. The significance of the Wald statistic in the column with heading Sig indicates the importance of the predictor variables in the model and high values of the Wald statistic shows that the corresponding predictor variable is significant.

$\exp\left(\hat{\beta}\right) = e^{\hat{\beta}}$, which is given in the last column of Table 4.5 is called the Odds Multiplier or odds ratio with an odds ratio of one for the reference category, and all other groups are compared on the basis of the reference group. Values greater than one indicate that the variable in question increases the odds of being HIV+ individual and values between 0 and 1 indicate a decrease in the odds of being HIV+ individual.

$\exp\left(\hat{\beta}\right)$, is a factor by which the odds of being HIV+ individual change when i^{th} independent variable increases by a unit. For instance from Table 4.5 the value of the odd ratio $\exp\left(\hat{\beta}\right) = 1.997$, for age indicates that clients aged less 15 years have nearly 2 times chances to be HIV+ than those in the reference category (50 years and above), clients aged between 15 and 30 years have 1.8 times chances to be HIV+ than those in the reference category and clients aged between 31 and 49 years have more than double chances to be HIV+ than those in the reference category and its effect is statistically significant.

Prior to using the fitted model for the intended purpose, it is necessary to assess and diagnose the goodness of fit of the model.

4.4 Assessing the Goodness of fit of the model

After fitting a model to a set of data, it is natural to enquire about the extent to which the fitted values of the response variable under the model compare with the observed values. If the agreement between the observations and the corresponding fitted values is good, the model may be acceptable. If not, the current form of model will certainly not be acceptable and the model will need to be revised. This aspect of the adequacy of a model is widely referred to as **goodness of fit**.

4.4.1 The likelihood ratio test

The likelihood ratio test of a model tests the difference between -2LL for the full model and -2LL for the initial chi-square in the null model. This is called the model chi-square test. The null model, also called the initial model, is logit(p)= the constant. That is, initial chi-square is -2LL for the model, which accepts the null hypothesis that all the β coefficients are 0. This implies that none of the independent variable is linearly related to the log odds of the dependent. The model chi-square thus tests the null hypothesis that all population logistic regression coefficients except the constant are zero. The degrees of freedom in this test equal the number of terms in the model minus 1 (for the constant). This is the same as the difference in the number of terms between the two models, since the null model has only one term. Model chi-square measures the improvement in the fit that the explanatory variables make compared to the null model. Model chi –square is a likelihood ratio test, which reflects the difference between the error of not knowing the independent variables (intial chi-square) and error when the independent variables are included in the model (deviance).

When P (model chi-square) ≤ 0.05, we reject the null hypothesis. As can be seen from Table 4.6, we reject the null hypothesis.

Table 4.6 Model fitting information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	4746.363			
Final	3973.203	773.161	12	.000

4.4.2 **Classification table**

An intuitively appealing way to summarize the results of a fitted logistic model is via a classification table. This table is the result of cross-classifying the outcome variable, Y, with a dichotomous variable whose values are derived from the estimated logistic probabilities (Hosmer and Lemeshow, 1989). A classification table displays the tabulated cross-classifications of the observed category by the model-predicted category on the dependent variable and it is one way of assessing the goodness of fit of a model. A good model is one that minimizes miss-classifications. As shown in Table 4.7 the fitted model has an overall predictive accuracy of 70.8% with 70.8% of the negative group, and 70.7% of the positive group being correctly classified and this may be considered sufficient.

Table 4.7 Classification Table for Y

Observed	Predicted		
	HIV status of an individual		Percentage Correct
	Negative	Positive	
Negative	9762	4018	70.8
Positive	275	663	70.7
Overall Percentage			70.8

4.4.3 The Hosmer-Lemeshow test

The hypothesis to be tested is:

H_0 : Model fits the data versus H_A :Model does not fit the data

Table 4.8: Hosmer-Lemeshow Test

Chi-square	Df	Sign.level
6.860	8	0.552

This shows that there is no sufficient evidence to reject the null hypothesis and it confirms that our model has a good fit.

By using the model chi-square goodness fit test, the classification table and the Hosmer-Lemeshow test we can say the fitted model is statistically satisfactory.

Therefore, the best fit of the data is given by

$$Logit\left(\frac{p_i}{1-p_i}\right) = -2.105 + \hat{\beta}_a X_{1i} + \hat{\beta}_s X_{2i} + \hat{\beta}_m X_{3i} + \hat{\beta}_e X_{4i} + \hat{\beta}_o X_{6i} + \hat{\beta}_{mi} X_{7i} +$$

$$\hat{\beta}_{mi} X_{7i} + \hat{\beta}_p X_{8i} + \hat{\beta}_{EH} X_{9i} + \hat{\beta}_{Se} X_{10i} + \hat{\beta}_{STI} X_{13i}$$

where p_i = the probability that the ith individual will get HIV under the given level of the explanatory variables, $\hat{\beta}_a$ =coefficient of the ath category of age, X_{1i} =age of the ith individual, $\hat{\beta}_s$ = coefficient of the sth category of sex, X_{2i} = sex of the ith individual, $\hat{\beta}_m$ =coefficient of the mth category of marital status, X_{3i} =marital status of the ith individual, $\hat{\beta}_e$ =coefficient of the eth category of educational level, X_{4i} =educational level of the ith individual, $\hat{\beta}_o$ =coefficient of the oth category of occupation, X_{6i} = occupation of the ith individual, $\hat{\beta}_{mi}$ =coefficient of the mith category of the most important reason why client is

here today, X_{7i} = the most important reason why client is here today of the ith individual, $\hat{\beta}_p$ = coefficient of the pth category of previously tested, X_{8i} = previously tested of the ith individual, $\hat{\beta}_{EH}$ = coefficient of the EH^{th} category of ever had sex with penetration, X_{9i} = ever-had sex with penetration of the ith individual, $\hat{\beta}_{se}$ = coefficient of the se^{th} category of suspected exposure time of the ith individual, X_{10i} = suspected exposure time of the ith individual, $\hat{\beta}_{STI}$ = coefficient of the STH^{th} category of history of STI of the ith individual, X_{13i} = history of STI of the ith individual. The estimated coefficients are given in Table 4.5 above.

4.5 Case Diagnostics

In addition to global examination of a model, it is also useful to examine the characteristics of individual cases in our data set. We are concerned with cases that might unduly influence our parameter estimates. Cook's distance is one way of analyzing influence statistic in logistic regression. Cook's distance is computed using SPSS as a measure of the influence, which a case has on the solution. This is the same statistic that is used as measure of influence in multiple regression. However, the criteria for determining that a case is influential in logistic regression differ from the criteria in multiple regression. In logistic regression, a case is identified as influential if its Cook's distance is greater than one. This is based on a statement in Hosmer and Lemeshow (1989): " In our experience the influence diagnostic must be larger than 1.0 for an individual covariate pattern to have an effected on the estimated coefficients." Since Cook's distances for each one of the observations Table 4.9 in our data is less than one we conclude that there are no influential observations.

Table 4.9 Cook's influence statistics

Mean	0.017581
Mode	0.00000
Minimum	0.00000
Maximum	0.85697

4.6 Interpretation of the logistic regression model

The logistic regression model indicates that HIV status of an individual (negative, positive) is affected by many factors considered in the study. The results given above for the logistic regression model in Table 4.5 can be used for discussion and interpretation. In our analyses of the prevalence of HIV infection, all the variables included in the model were found to have significant effect at $p<0.05$. From the fitted logistic regression model, when compared with individuals in the age group ≥ 50 (Ref), the odds of HIV + individuals was found to be highest (above twice the reference category) for individuals in the age group 31-49 followed by those below 15 years (almost twice the reference category) and 15-30 years (1.8 times the reference category). From this multivariate finding it can be inferred that the prevalence of HIV infection is high for those in the age group 31-49 years in urban Addis Abeba. The coefficients in Table 4.5 also indicate that the prevalence of HIV infection varies by sex. The odds of being HIV+ for male clients is about 50.2% less likely to be infected with HIV than female (ref) clients and its effect is statistically significant. This indicates that the problem is severe for females than males in urban Addis Abeba. The result for marital status indicate that the never-married, married and divorced clients are 25.7%, 27.5% and 50.5% less likely to be infected with HIV respectively than widowed clients, implying clients whose marriage ended because of death of a partner is found to have a significantly high likelihood of being exposed to HIV infection in urban Addis Abeba. The level of education is grouped into four categories ranging from those who are illiterate to those who have attended tertiary level education. The odds of being HIV + individual with education level illiterate, grade (1-6) and junior and high

school was found to be 1.9, 1.5 and 1.6 times that of above grade 12 respectively, implying that those with higher educational attainment (tertiary) exhibited a lower chance to be HIV+ as compared to illiterate. Occupation of a client is one of the determinant of HIV status of an individual. Students are about 31% less likely to be infected with HIV than those in the reference category the reference category is other which includes housewives, armed and police forces,...etc. Clients who never had sex are about 50.9% less likely to be infected with HIV than those clients who had started sex and its effect is statistically significant. The large positive model odds ratio (1.648) for previously tested indicates that the prevalence of HIV infection is higher among clients who had not previously tested than those previously tested. Clients who had not previously tested are about 64.8% likely to be infected with HIV than those who had. Regarding suspected exposure time of a client to HIV, it is found that individuals whose exposure time is less than one month and over 6 months are twice likely to be infected than those clients who are not exposed (ref). Clients whose exposure time is between 1 to 3 months are 1.6 times likely to be infected than clients who are not exposed (ref). Clients whose exposure time is between 4 to 6 months are 93.6% less likely to be infected than clients who are not exposed. A lifetime history of STI has also a large and statistically significant model odds ratio (2.02), indicating that the high level of HIV infection is exhibited among clients that have a history of STI. The results also indicate that the most important reason clients visit the OSSA mobile VCT center determines the HIV status of an individual. Clients who visited the center due to their risky behavior they passed through are about 71.1% less likely to be infected with HIV than other which is the reference category that includes clients that visit the center due to illness, marital reunion, test before pregnancy ...etc. Clients who visit the center for the reason that they are visa applicants were found to have 57.4% less likely to be infected than the reference category, clients who visit the center because they want to plan for future are 34% less likely to be infected than the reference category.

CHAPTER FIVE

Conclusions and Recommendations

5.1 Conclusions

This study is an attempt to examine the impact of some pre-test counselling session factors such as demographic and risk behavior characteristic of VCT clients at OSSA mobile VCT center in urban Addis Ababa.

The socio-demographic profile of VCT clients at the OSSA mobile VCT center in Addis Abeba showed the age 31-49 years to be the most affected age group by HIV/AIDS. Targeting this age group would greatly reduce the risk of HIV/AIDS. Female clients are more likely to be HIV positive than males (4.5% vs. 9.7%). This pattern of female vulnerability to HIV/AIDS infection in Addis Ababa is commonly seen in many other studies. Perhaps due to the biological factors as well as the prevailing low socio economic status of women and other cultural influences, targeting women for HIV prevention and control is mandatory. It was also observed that those clients whose marriage doom ended by death of a partner or widowed clients have higher prevalence of HIV infection. HIV infection has no association with one's employment status. This indicates that the level of expansion of HIV is the same for both active and inactive clients. Educational levels of clients have a significant influence on HIV infection, as the level of education increases the prevalence of HIV decreases. As vividly shown in the analyses, the clients that are at a higher risk group of HIV infection and this was supported by a majority of those who mentioned having risky behavior as the main reason for seeking VCT services. It was also observed that those clients who had sexual practices in the past are liable for HIV/AIDS. This clearly shows that sexual intercourse is the most important factor that puts people at risk of contracting HIV/AIDS. The study further indicated that condom use was low among VCT clients, with only 15.4% of clients reporting using condoms in three months prior to testing. Female are unlikely to use condoms than their male counterparts. This difference could be attributed to women low negotiation skill

on the use of condom or men's negligence and engagement in high-risk sexual behavior. Furthermore the findings showed that a lifetime history of STI has a statistically significant impact on the status of HIV/AIDS of an individual.

In conclusion, low socioeconomic status and high-risk behavior are found to be significantly associated with HIV infection in urban Addis Ababa.

5.2 Recommendations

Based on our interpretation of the results of this study, we offer the following recommendations:

- ✓ Young women are especially vulnerable to HIV/AIDS. In order to address this problem, continuous awareness raising programs should be devised so as to bring about behavioral change that will in turn minimize the risk of being vulnerable to HIV/AIDS.
- ✓ For those who engaged in unsafe sexual practices the perception that hinders people from not using condoms need to be properly addressed.
- ✓ Address religious and cultural barriers to gender equality.
- ✓ VCT program designers develop messages and modes of communication that will attract youth who are at high risk of contracting HIV/AIDS.
- ✓ This study is done only for OSSA mobile VCT center in urban Addis Ababa based on only some factors and cannot be taken as conclusive. Hence, we further recommend similar studies with additional exogenous variables (such as religion, location of residence, ethnicity...etc.) for other big towns in the country and for the country as a whole so as to provide complete and useful insight for formulating appropriate policies. Hence it is recommended for further studies in the area.

REFERENCES

1. **Agresti, A (1996).** An Introduction to categorical data Analysis. John Wiley and Sons, Inc., New York.
2. **Amuyunzu-Myamongo M, Biddlecom AE, Ouedraogo C, Woog V (2005).** Qualitative evidence on adolescent's views of sexual and reproductive health in Sub-Saharan Africa. The Alan Guttmacher Institute. Report No: 1, New York.
3. **Antenane Kora, Mebiratu Bejiga and Solomon Tesfaye (2005).** Socio-demographic profile and prevalence of HIV infection among VCT clients in Addis Ababa, Ethiop. J. Health Dev, 19(2): 109 – 116.
4. **Bekele Abebe,** Factors influencing the use of voluntary counselling and testing services in poor communities of Addis Ababa. 2000 U.S. CDC, Addis Ababa, Ethiopia
5. **Bradley Heather, Amare Bedada, Amy Tsui, Heena Brahmbhtt, Duff Gillespie, Aklilu Kidanu.** HIV and family planning service integration and voluntary HIV counselling and testing client's composition in Ethiopia 2007. Addis Ababa, AIDS and Behavior.
6. **Cartoux M, Meda N, Van de Perre P, Newell ML, de Vincenti I, Dabis F.** Acceptability of voluntary HIV testing by pregnant women in developing countries an international survey. Ghent international working group on mother-to-child transmission of HIV. AIDS 1998:12; 489-93, Geneva.
7. **Central Statistical Authority (CSA) and ORC Macro.** Demographic and Health survey of Ethiopia. CSA, Addis Ababa, 2005:213-221.
8. **Central Statistical Authority (CSA), Ethiopia (2004).** Statistical Abstract.
9. **Coates TJ, Grinstead OA, Gregorich SE, Sweat MD, Kamenga MC, Sangiwa G.** Efficacy of Voluntary HIV counselling and testing in individuals and couples in Kenya, Tanzania and Trindad: A randomized trial. The Voluntary HIV Counselling and Testing Efficacy Study Group. Lancet 2000; 356: 103-13

10. **Collett, D. (1991).** Modeling Binary Data. Chapman and Hall, London.

11. **Corey L and Adamson S.** HIV/AIDS in Sub-Saharan Africa Croatian Medical Journal 2004, 45(4): 402-414,

12. **Depaoli MM, Mangoni R, Kleep KI (2004).** Factors influencing acceptability of voluntary counselling and HIV-testing among pregnant women in northern Tanzania. AIDS Care 2004; 16: 411-25

13. **Allen E and Glynn E.** Prevention of MTCT of HIV in Africa. Practical guidance for programs. June 2001

14. **Family Guidance Association of Ethiopia.** A study on factors affecting accessibility and acceptability of VCT service for HIV/AIDS in Bahirdar town, 2001.

15. **Federal Ministry of Health/ National HIV/AIDS Prevention and Control Office.** AIDS in Ethiopia 6 th Report, 2006.

16. **Fylkesness K and Siziya S.** A randomized trial on acceptability of voluntary HIV counselling and testing. Trop Med Int Health; 9:566-72, 2004.

17. **Hosmer, D and Lemeshow, S (1989).** Applied Logistic Regression. John Wiley and Sons. Inc., New York.

18. **Houston S, Archibald CP, Strike C, Sutherland D.** Factors associated with HIV testing among Canadians: Results of a population-based survey. Int J STD AIDS 9: 341-6, 1998.

19. **Kim YM, Marangwanda C, Kols A.** Quality of counselling of young clients in Zimbabwe. East Afr Med J; 74: 514-8, 1997.

20. **Kipp W, Kabagambe G and Konde-lule J.** Low impact of a community-wide HIV testing and counselling program on sexual behavior in rural Uganda. Educ Prev 13: 279-89, 2001.

21. **Logistic Regression:** www.http// file:/PA 765 Logistic Regression.html.

22. **Main DS, Iverson DC and Mcgloin J.** Comparison of HIV-risk behavior and demographics of adolescents tested or not tested for HIV infection public Health Rep 109:699-702, 1994.

23. **Maman S, Mbwambo J, Hogan NM, Kilonto GP, Sweat M.** Women's barrier to HIV testing disclosure: Challenges for HIV voluntary counselling and testing. AIDS Care 13: 595-603, 2001.

24. **Matovu JK, Gray RH, Makumbi F, Wawer MJ. Serwedda D; Kigozi G.** Voluntary HIV counselling and testing acceptance, sexual risk behavior and HIV incidence in Rakai, Uganda. AIDS 19: 503-11,2005.

25. **Menard, S. (1995).** Applied Logistic Regression Analysis. Sage publication series: Quantitative Applications in the social sciences, No.106.

26. **Ministry of Health.** Health and Health Related Indicators. Addis- Ababa, Ethiopia, 2000.

27. **Mitike G, Wuleta L, Frehiwot B, Reta A, Tamrat S, Tewodros G, Fikre E, Atalay A, Yigremu A, and Dereje K.** HIV/AIDS behavioral surveillance survey (BSS) Ethiopia, round one, Ministry of Health, HAPCO, 2002.

28. **Muller O, Barugahare L, Schwartlander B, Byaruhanga E, Kataaha P, Kyeyune D.** HIV prevalence, attitudes and behaviors in clients of a confidential HIV testing and counselling center in Uganda. AIDS 6: 869 - 74, 1992.

29. **Nuaha F, Kabatesi D, Mugnawa M, Whalen CC.** Factors influencing acceptability of voluntary counselling and testing for HIV in Bushenyi district of Uganda. East Afr Med J 79: 626-32, 2002.

30. **Painter TM.** Voluntary counselling and testing for couples: a high-leverage intervention for HIV/AIDS prevention in Sub-Saharan Africa soc sci Med 53: 1397-411, 2001.

31. **Pool R, Nyanzi S and Whitworth JA.** Attitudes to voluntary counselling and testing for HIV among pregnant women in rural south west Uganda. AIDS Care 13: 605-15, 2001.

32. **Sahilu T, Tsegaye A, Rinke T, Agonafir T, and W/Michael T.** Effect of an intervention based on voluntary HIV testing and counselling on HIV incidence among ENRAP cohort study participants. A special issue on HIV/AIDS in Ethiopia. Ethiopian Medical Journal, 1999:37(1): 119.

33. Samet JH, Winter MR, Grant L and Hingson R. Factors associated with HIV testing among sexually active adolescents a Massachsetts survey. Pediatrics 100: 371-7, 1997.

34. UN AIDS " Voluntary Counselling and Testing " best practice collection, technical update Geneva UN/AIDS, May 2000.

35. UNAIDS. Report on the Global HIV/AIDS Epidemic. Geneva., 2006.

36. UNSECO. HIV/AIDS Impact on Education Clearinghouse, 2002. Available at website: http:// www.hivaidsclearinghouse.unesco.org

37. Van Dyk AC. " What is the point of knowing?" Psychosocial barriers to HIV/AIDS voluntary counselling and testing programmes in South-Africa. S. Afr J Psychol 33: 118-25, 2003.

38.**Weinhardt LS, Carey MP, Johnston BT and Bickham NL.** Effects of HIV counselling and testing on sexual risk behavior. A meta-analytic review of published research 1985-1997. AMJ Public Health, 89:1397-405, 1999.

APPENDIX

1. Contingency Table for Hosmer and Lemeshow Test

		HIV status of an indiviual = negative		HIV status of an indiviual = positive		
		Observed	Expected	Observed	Expected	Total
Step 1	1	8737	8737.000	384	384.000	9121
	2	4068	4068.000	306	306.000	4374
	3	975	975.000	248	248.000	1223
Step 2	1	3539	3530.454	52	60.546	3591
	2	1328	1336.234	43	34.766	1371
	3	1716	1713.940	68	70.060	1784
	4	996	998.233	62	59.767	1058
	5	3030	3042.507	239	226.493	3269
	6	544	548.286	54	49.714	598
	7	1722	1706.297	177	192.703	1899
	8	905	904.049	243	243.951	1148
Step 3	1	1698	1691.612	6	12.388	1704
	2	1688	1693.822	36	30.178	1724
	3	1170	1181.494	43	31.506	1213
	4	1546	1540.448	44	49.552	1590
	5	1353	1350.210	57	59.790	1410
	6	1512	1507.741	95	99.259	1607
	7	1237	1242.332	111	105.668	1348
	8	1158	1170.612	134	121.388	1292
	9	1631	1616.000	182	197.000	1813
	10	787	785.730	230	231.270	1017
Step 4	1	1705	1698.305	6	12.695	1711
	2	1401	1404.494	25	21.506	1426
	3	1752	1751.236	40	40.764	1792
	4	1403	1407.316	54	49.684	1457
	5	1479	1489.267	79	68.733	1558
	6	1523	1521.944	98	99.056	1621
	7	1662	1659.321	137	139.679	1799
	8	1330	1322.074	136	143.926	1466
	9	1525	1526.043	363	361.957	1888
Step 5	1	1397	1393.476	5	8.524	1402
	2	1458	1458.600	19	18.400	1477
	3	1433	1434.544	30	28.456	1463
	4	1407	1412.785	44	38.215	1451
	5	1374	1369.350	49	53.650	1423
	6	1428	1438.912	84	73.088	1512
	7	1414	1413.523	99	99.477	1513
	8	1429	1437.857	145	136.143	1574
	9	1346	1319.645	140	166.355	1486
	10	1094	1101.309	323	315.691	1417
Step 6	1	1396	1391.507	4	8.493	1400
	2	1253	1255.606	17	14.394	1270
	3	1454	1459.017	31	25.983	1485
	4	1361	1358.411	32	34.589	1393
	5	1458	1456.246	51	52.754	1509

	6	1401	1406.454	71	65.546	1472
	7	1383	1382.156	93	93.844	1476
	8	1375	1377.658	122	119.342	1497
	9	1326	1320.005	147	152.995	1473
	10	1373	1372.941	370	370.059	1743
Step 7	1	1556	1552.159	6	9.841	1562
	2	1491	1492.061	19	17.939	1510
	3	1433	1438.254	33	27.746	1466
	4	1446	1445.839	39	39.161	1485
	5	1478	1474.222	53	56.778	1531
	6	1430	1429.716	75	75.284	1505
	7	1435	1433.760	102	103.240	1537
	8	1327	1346.973	150	130.027	1477
	9	1319	1310.087	181	189.913	1500
	10	865	856.929	280	288.071	1145
Step 8	1	1681	1679.171	9	10.829	1690
	2	1475	1472.182	15	17.818	1490
	3	1429	1434.240	33	27.760	1462
	4	1435	1432.043	36	38.957	1471
	5	1450	1452.228	58	55.772	1508
	6	1399	1397.662	72	73.338	1471
	7	1384	1376.330	89	96.670	1473
	8	1323	1338.999	144	128.001	1467
	9	1295	1287.819	176	183.181	1471
	10	909	909.327	306	305.673	1215
Step 9	1	1428	1424.034	3	6.966	1431
	2	1454	1457.181	18	14.819	1472
	3	1443	1447.242	29	24.758	1472
	4	1440	1437.876	36	38.124	1476
	5	1442	1440.505	50	51.495	1492
	6	1400	1395.320	66	70.680	1466
	7	1491	1492.174	103	101.826	1594
	8	1347	1356.734	137	127.266	1484
	9	1305	1298.136	167	173.864	1472
	10	1030	1030.797	329	328.203	1359
Step 10	1	1389	1387.589	5	6.411	1394
	2	1514	1513.955	15	15.045	1529
	3	1478	1481.542	29	25.458	1507
	4	1429	1437.430	46	37.570	1475
	5	1431	1417.432	38	51.568	1469
	6	1401	1402.414	72	70.586	1473
	7	1379	1375.841	90	93.159	1469
	8	1369	1367.965	125	126.035	1494
	9	1295	1301.174	177	170.826	1472
	10	1095	1094.657	341	341.343	1436

2. Variables not in the Equation

			Score	df	Sig.
Step 1	Variables	x1	28.804	3	.000
		x1(1)	.000	1	.985
		x1(2)	2.871	1	.090
		x1(3)	19.756	1	.000
		x2(1)	49.834	1	.000
		x4	60.231	3	.000
		x4(1)	17.776	1	.000
		x4(2)	5.304	1	.021
		x4(3)	.292	1	.589
		x6	121.829	4	.000
		x6(1)	86.406	1	.000
		x6(2)	2.259	1	.133
		x6(3)	16.063	1	.000
		x6(4)	.465	1	.495
		rex7	165.571	3	.000
		rex7(1)	129.578	1	.000
		rex7(2)	.037	1	.847
		rex7(3)	165.002	1	.000
		x8(1)	37.723	1	.000
		x9(1)	108.079	1	.000
		x10	218.542	4	.000
		x10(1)	3.745	1	.053
		x10(2)	.002	1	.966
		x10(3)	6.907	1	.009
		x10(4)	160.666	1	.000
		x11	77.649	3	.000
		x11(1)	35.117	1	.000
		x11(2)	1.772	1	.183
		x11(3)	10.077	1	.002
		x12	109.271	3	.000
		x12(1)	57.587	1	.000
		x12(2)	.214	1	.644
		x12(3)	.370	1	.543
		rex13(1)	57.525	1	.000
	Overall Statistics		582.326	27	.000
Step 2	Variables	x1	25.131	3	.000
		x1(1)	.013	1	.909
		x1(2)	2.422	1	.120
		x1(3)	16.472	1	.000
		x2(1)	66.928	1	.000
		x4	55.847	3	.000
		x4(1)	16.190	1	.000
		x4(2)	4.980	1	.026
		x4(3)	.229	1	.632
		x6	90.785	4	.000
		x6(1)	48.264	1	.000
		x6(2)	.118	1	.732
		x6(3)	20.730	1	.000
		x6(4)	.226	1	.635
		rex7	65.198	3	.000
		rex7(1)	38.326	1	.000
		rex7(2)	.192	1	.661
		rex7(3)	63.492	1	.000
		x8(1)	37.349	1	.000
		x9(1)	14.524	1	.000
		x11	12.073	3	.007
		x11(1)	4.830	1	.028
		x11(2)	11.141	1	.001
		x11(3)	1.260	1	.262
		x12	19.474	3	.000
		x12(1)	13.610	1	.000
		x12(2)	4.737	1	.030
		x12(3)	.025	1	.874
		rex13(1)	39.856	1	.000
	Overall Statistics		364.752	23	.000

			Score	df	Sig.
Step 3	Variables				
			23.841	3	.000
		x1			
		x1(1)	.015	1	.902
		x1(2)	.611	1	.434
		x1(3)	12.164	1	.000
		x2(1)	70.455	1	.000
		x4	19.386	3	.000
		x4(1)	8.832	1	.003
		x4(2)	.444	1	.505
		x4(3)	.253	1	.615
		rex7	73.503	3	.000
		rex7(1)	43.009	1	.000
		rex7(2)	.124	1	.725
		rex7(3)	71.531	1	.000
		x8(1)	36.633	1	.000
		x9(1)	4.315	1	.038
		x11	10.732	3	.013
		x11(1)	2.711	1	.100
		x11(2)	10.635	1	.001
		x11(3)	1.071	1	.301
		x12	8.594	3	.035
		x12(1)	7.801	1	.005
		x12(2)	4.498	1	.034
		x12(3)	.004	1	.953
		rex13(1)	29.482	1	.000
	Overall Statistics		275.507	19	.000
Step 4	Variables	x1	25.242	3	.000
		x1(1)	.060	1	.806
		x1(2)	6.307	1	.012
		x1(3)	20.884	1	.000
		x4	11.038	3	.012
		x4(1)	2.620	1	.106
		x4(2)	.180	1	.671
		x4(3)	.048	1	.826
		rex7	72.860	3	.000
		rex7(1)	46.878	1	.000
		rex7(2)	.013	1	.911
		rex7(3)	72.077	1	.000
		x8(1)	42.359	1	.000
		x9(1)	6.221	1	.013
		x11	6.782	3	.079
		x11(1)	.277	1	.599
		x11(2)	5.339	1	.021
		x11(3)	2.836	1	.092
		x12	5.340	3	.149
		x12(1)	3.138	1	.076
		x12(2)	.806	1	.369
		x12(3)	.005	1	.942
		rex13(1)	46.981	1	.000
	Overall Statistics		210.376	18	.000
Step 5	Variables	x1	27.263	3	.000
		x1(1)	.016	1	.901
		x1(2)	8.588	1	.003
		x1(3)	23.859	1	.000
		x4	13.071	3	.004
		x4(1)	3.050	1	.081
		x4(2)	.154	1	.695
		x4(3)	.125	1	.724
		x8(1)	37.294	1	.000
		x9(1)	9.518	1	.002
		x11	5.331	3	.149
		x11(1)	.564	1	.453
		x11(2)	3.390	1	.066
		x11(3)	2.506	1	.113
		x12	8.314	3	.040
		x12(1)	3.502	1	.061
		x12(2)	.411	1	.521
		x12(3)	.003	1	.960
		rex13(1)	52.559	1	.000
	Overall Statistics		141.538	15	.000

Step 6	Variables	x1	25.824	3	.000
		x1(1)	.002	1	.963
		x1(2)	4.041	1	.044
		x1(3)	19.057	1	.000
		x4	12.025	3	.007
		x4(1)	3.158	1	.076
		x4(2)	.002	1	.967
		x4(3)	.304	1	.581
		x8(1)	32.360	1	.000
		x9(1)	9.059	1	.003
		x11	3.973	3	.264
		x11(1)	.574	1	.449
		x11(2)	2.801	1	.094
		x11(3)	1.562	1	.211
		x12	8.355	3	.039
		x12(1)	3.829	1	.050
		x12(2)	.515	1	.473
		x12(3)	.122	1	.727
	Overall Statistics		88.320	14	.000
Step 7	Variables	x1	27.907	3	.000
		x1(1)	.011	1	.915
		x1(2)	2.596	1	.107
		x1(3)	18.377	1	.000
		x4	9.351	3	.025
		x4(1)	1.630	1	.202
		x4(2)	.384	1	.536
		x4(3)	1.550	1	.213
		x9(1)	12.476	1	.000
		x11	3.973	3	.264
		x11(1)	.458	1	.498
		x11(2)	1.908	1	.167
		x11(3)	1.913	1	.167
		x12	11.174	3	.011
		x12(1)	3.781	1	.052
		x12(2)	.168	1	.682
		x12(3)	.138	1	.711
	Overall Statistics		55.688	13	.000
Step 8	Variables	x4	10.927	3	.012
		x4(1)	2.923	1	.087
		x4(2)	.095	1	.758
		x4(3)	.657	1	.418
		x9(1)	11.517	1	.001
		x11	3.649	3	.302
		x11(1)	.639	1	.424
		x11(2)	2.074	1	.150
		x11(3)	1.431	1	.232
		x12	10.443	3	.015
		x12(1)	3.902	1	.048
		x12(2)	.269	1	.604
		x12(3)	.162	1	.688
	Overall Statistics		27.787	10	.002
Step 9	Variables	x4	10.726	3	.013
		x4(1)	2.821	1	.093
		x4(2)	.106	1	.745
		x4(3)	.687	1	.407
		x11	4.180	3	.243
		x11(1)	.045	1	.833
		x11(2)	3.207	1	.073
		x11(3)	1.023	1	.312
		x12	1.540	3	.673
		x12(1)	1.364	1	.243
		x12(2)	1.007	1	.316
		x12(3)	.201	1	.654
	Overall Statistics		16.104	9	.065
Step 10	Variables	x11	4.197	3	.241
		x11(1)	.007	1	.932
		x11(2)	2.992	1	.084
		x11(3)	1.109	1	.292
		x12	1.379	3	.710
		x12(1)	1.222	1	.269
		x12(2)	.896	1	.344
		x12(3)	.168	1	.682
	Overall Statistics		5.394	6	.494

3. Model if Term Removed

Variable		Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of the Change
Step 1	x3	-3489.840	365.694	3	.000
Step 2	x3	-3307.258	240.524	3	.000
	x10	-3306.993	239.993	4	.000
Step 3	x3	-3229.199	186.048	3	.000
	x6	-3186.996	101.641	4	.000
	x10	-3235.652	198.953	4	.000
Step 4	x2	-3136.175	67.724	1	.000
	x3	-3161.875	119.124	3	.000
	x6	-3154.830	105.033	4	.000
	x10	-3205.507	206.388	4	.000
Step 5	x2	-3097.466	67.334	1	.000
	x3	-3125.149	122.699	3	.000
	x6	-3121.029	114.460	4	.000
	rex7	-3102.313	77.028	3	.000
	x10	-3116.462	105.326	4	.000
Step 6	x2	-3081.696	81.965	1	.000
	x3	-3095.300	109.171	3	.000
	x6	-3092.311	103.193	4	.000
	rex7	-3081.547	81.666	3	.000
	x10	-3090.337	99.247	4	.000
	rex13	-3063.799	46.170	1	.000
Step 7	x2	-3066.677	85.631	1	.000
	x3	-3074.328	100.932	3	.000
	x6	-3075.614	103.505	4	.000
	rex7	-3062.243	76.764	3	.000
	x8	-3040.714	33.705	1	.000
	x10	-3073.133	98.543	4	.000
	rex13	-3044.320	40.918	1	.000
Step 8	x1	-3023.862	29.874	3	.000
	x2	-3049.550	81.251	1	.000
	x3	-3054.998	92.146	3	.000
	x6	-3058.868	99.886	4	.000
	rex7	-3047.581	77.313	3	.000
	x8	-3027.122	36.394	1	.000
	x10	-3057.189	96.527	4	.000
	rex13	-3029.482	41.115	1	.000
Step 9	x1	-3017.297	28.844	3	.000
	x2	-3044.488	83.225	1	.000
	x3	-3048.295	90.840	3	.000
	x6	-3044.951	84.152	4	.000
	rex7	-3043.578	81.405	3	.000
	x8	-3022.964	40.178	1	.000

	x9	-3008.925	12.099	1	.001
	x10	-3030.442	55.134	4	.000
	rex13	-3023.297	40.844	1	.000
Step 10	x1	-3012.493	30.372	3	.000
	x2	-3034.401	74.187	1	.000
	x3	-3040.797	86.980	3	.000
	x4	-3002.875	11.136	3	.011
	x6	-3025.630	56.647	4	.000
	rex7	-3038.930	83.245	3	.000
	x8	-3015.337	36.061	1	.000
	x9	-3003.251	11.888	1	.001
	x10	-3024.488	54.362	4	.000
	rex13	-3017.386	40.158	1	.000

Collinearity Diagnostics

Model Dimension	Eigenvalue	Condition index	Variance Proportions												
			(constant)	Age in year	Sex	Marital status	Educational level	Occupation	The most important reason why client is here today	Previously tested	Ever had sex with penetration	Suspected exposure time	Condom use last 3 months	Used condom last time had sex	History of STI
1	8.413	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	1.632	2.27	0.00	0.00	0.00	0.04	0.00	0.01	0.00	0.01	0.00	0.00	0.03	0.02	0.00
3	0.723	3.412	0.00	0.00	0.43	0.09	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00
4	0.641	3.623	0.00	0.1	0.3	0.02	0.00	0.01	0.02	0.46	0.00	0.00	0.00	0.00	0.00
5	0.452	4.314	0.00	0.00	0.07	0.14	0.01	0.00	0.64	0.00	0.01	0.00	0.00	0.00	0.00
6	0.404	4.566	0.00	0.00	0.07	0.42	0.00	0.04	0.22	0.24	0.00	0.00	0.02	0.01	0.00
7	0.250	5.796	0.00	0.00	0.01	0.00	0.06	0.76	0.04	0.03	0.00	0.00	0.02	0.01	0.01
8	0.134	7.916	0.00	0.03	0.00	0.04	0.01	0.00	0.06	0.01	0.05	0.40	0.34	0.02	0.00
9	0.110	8.727	0.00	0.15	0.00	0.13	0.52	0.13	0.00	0.02	0.03	0.05	0.17	0.05	0.00
10	0.096	9.374	0.00	0.54	0.08	0.11	0.11	0.00	0.00	0.01	0.01	0.12	0.00	0.01	0.17
11	0.084	10.015	0.00	0.05	0.01	0.00	0.18	0.01	0.00	0.01	0.00	0.28	0.30	0.27	0.18
12	0.050	12.983	0.01	0.14	0.00	0.00	0.03	0.02	0.01	0.00	0.38	0.02	0.11	0.38	0.39
13	0.011	27.589	0.99	0.08	0.03	0.00	0.08	0.02	0.01	0.00	0.51	0.14	0.00	0.22	0.24

DECLARATION

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any other University and all sources of material used for the thesis have been duly acknowledged

Name: Hussein Mohammed

Signature: …………………………

Place: Faculty of Science, Addis Ababa University

Date: July, 2007

This thesis has been submitted for examination with my approval as a University advisor.

……………………………………………………………..

Professor Eshetu Wencheko